REMARKS

The Applicants correct a minor grammatical error in the specification.

In the outstanding Office Action, the Examiner rejects claims 1-28 as being anticipated by Jain et al., U.S. Patent No. 5,915,250. Applicants respectfully disagree.

Claim 1 recites the following: "deriving a set of perceptual semantic categories for representing important semantic cues in the human perception of images, where each semantic category is modeled through a combination of perceptual features that define the semantics of that category and that discriminate that category from other categories, wherein ***the perceptual features and their combinations*** are derived at least in part through ***subjective experiments*** performed with human observers" (emphasis added). Original, unamended independent claim 25 also recites similar subject matter: "program instructions for processing a set of perceptual semantic categories for representing semantic cues related to the manner in which human observers perceive and organize images, the semantic categories ... comprising a combination of perceptual features that define the semantics of a particular category and that discriminate that category from other categories, where ***the perceptual features and their combinations*** are derived through ***subjective experiments*** performed with human observers" (emphasis added).

Similar subject matter is also found in previously amended claim 14: "said data processor operating in accordance with a stored program for determining the semantic meaning of images in accordance with ***a set of perceptual semantic categories*** that were previously derived at least in part through ***subjective experiments*** performed with human observers and that represent important semantic cues in the human perception of images" (emphasis added). In this manner, amended claim 14 reduces issues for appeal and parallels the subject matter in claims 1 and 25.

Applicants describe subjective experiments (and the determination of semantic meaning, perceptual features and their combinations or perceptual semantic categories), e.g., as follows:

> ***A series of experiments were conducted***: 1) ***an image similarity experiment*** aimed at developing and refining a set of perceptual categories for photographic image databases, 2) ***a category naming and description experiment*** aimed at deriving a semantic name for each category, and a set of low-level features which describe it; and 3) ***an image categorization experiment*** to test the results of the metric, derived from the previous experiments, against the judgments of human observers on a new set of photographic images.
>
> All of the images in these experiments were selected from standard CD image collections, and provided high image quality and broad content. The images were selected according to the following criteria. First, a wide range of topics was included: people, nature, buildings, texture, objects, indoor scenes, animals, etc. Following a book designed to teach photography, the images were explicitly selected to include equal proportions of wide-angle, normal, and close-up shots, in both landscape and portrait modes. The selection of images was iterated so that it included images with different levels of brightness and uniform color distribution. Three sets of images (Set 1, Set 2 and Set3) included 97 images, 99 images and 78 images, respectively. The size of each printed image was approximately 1.5×1 inches (for a landscape), or 1×1.5 inches (for a portrait). All images were printed on white paper using a high-quality color printer.
>
> ***Seventeen subjects participated in these experiments ranging in age from 24 to 65. All of the subjects had normal or corrected-to-normal vision and normal color vision. The subjects were not familiar with the input images.***

Page 14, lines 4-24 (emphasis added). A particular exemplary experiment is described in part as follows:

> A purpose to ***a first experiment***, Experiment 1: Similarity Judgments for Image Set 2 to derive the Final Set of Semantic Categories, was to collect a second set of similarity judgments which enabled: 1) examining the perceptual validity and reliability of the categories identified by the hierarchical cluster analysis, 2) developing a final set of categories based on the similarity data for Set 1 and Set 2, and 3) establishing the connections between the categories.
>
> For this experiment, 97 thumbnails of all the images in Set 1 were printed, organized by cluster, and fixed to a tabletop, according to their initial categories, IC. The images were organized with a clear spatial gap between the different categories. Also printed were thumbnails of images from

> Set 2 (the new set). ***Twelve subjects (7 male and 5 female) participated in this experiment. Subjects were asked to assign each image from Set 2 into one of the initial categories, placing them onto the tabletop so that the most similar images were near each other.*** No instructions were given concerning the characteristics on which the similarity judgments were to be made, since this was the very information that the experiment was designed to uncover.. The order of the stimuli in Set 2 was random and different for each subject. This was done to counterbalance any effect the ordering of the stimuli might have on the subjective judgments. ***The subjects were not allowed to change the initial categories, as these images were fixed to the tabletop and could not be moved. However, subjects were allowed to do whatever they wished with the new images. They were free to change their assignments during the experiment, move images from one category into another, keep them on the side and decide later, or to start their own categories. Finally, at the end of the experiment, the subjects were asked to explain some of their decisions (as will be described later, these explanations, as well as the relative placement of images within the categories, were valuable in data analysis).***

Page 15, line 16 to page 16, line 7 (emphasis added).

The Applicants then describe, from page 16, line 9 to page 17, line 22, data analysis performed, e.g., to determine perceptual semantic categories. The Applicants then state the following:

> After the final categories had been identified, ***another experiment was performed*** to determine whether ***these algorithmically-derived categories were semantically distinct***. In this experiment, ***observers were requested*** to give names to the final categories identified in the first experiment. To further delineate the categories, and to identify high-level image features that discriminate the categories perceptually, ***the observers were also requested*** to provide descriptors for each of the categories. ***Each subject was asked*** to name each category and to write a brief description and main properties of the category. This experiment was helpful in many different ways. First, it was used to test the robustness of the categories and test whether people see them in a consistent manner. Furthermore, the experiment helped in establishing if the determined categories are semantically relevant. And finally, the written explanations are valuable in determining pictorial features that best capture the semantics of each category.
>
> A non-exhaustive listing of categories and their semantics are as follows.

C1: Portraits and close-ups of people. A common attribute for all images in this group is a dominant human face.

C2a: People outdoors. Images of people, mainly taken outdoors from medium viewing distance.

C2b: People indoors. Images of people, mainly taken indoors from medium viewing distance.

C3: Outdoor scenes with people. Images of people taken from large viewing distance. People are shown in the outdoor environment, and are quite small relative to image.

C4: Crowds of people. Images showing large groups of people on a complex background.

C5: Cityscapes. Images of urban life, with typical high spatial frequencies and strong angular patterns.

C6: Outdoor architecture. Images of buildings, bridges, architectural details that stand on their own (as opposed to being in a cityscape).

C7: Techno-scenes. Many subjects identified this category as a transition from C5 to C6.

C8a: Objects indoors. Images of man-made object indoors, as a central theme.

Page 17, line 24 to page 18, line 22.

These cited portions of the application therefore make it clear as to the types of exemplary subjective experiments that might be performed with human observers. Furthermore, the Applicants also detail in the cited text of the specification and at page 19 through page 25 additional aspects of deriving perceptual features and their combinations and perceptual semantic categories by using at least the subjective experiments.

Applicants will examine each cited portion of Jain to illustrate that such subjective experiments and the perceptual features and their combinations or perceptual semantic categories derived therefrom are not disclosed by (or implied by) Jain.

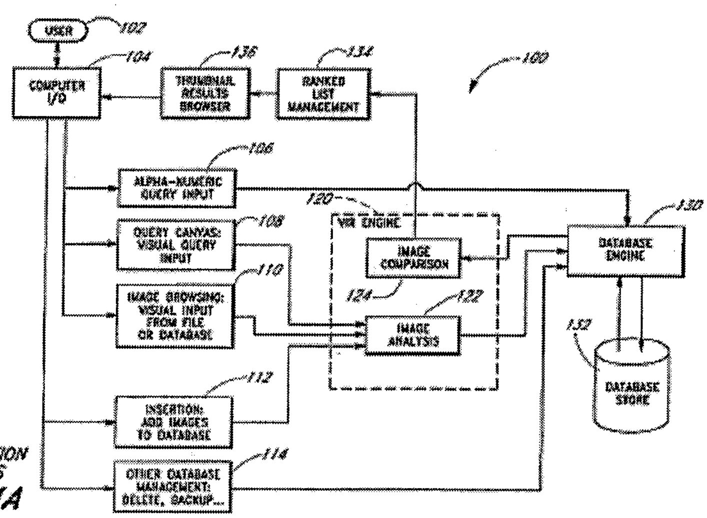
S.N. 10/033,597
Art Unit: 2624

The Examiner cites col. 4, lines 21-32 of Jain:

> Query by image property, wherein a user specifies a property or attribute of the image, such as the arrangement of colors, or they may sketch an object and request the system to find images that contain similar properties. The Engine also allows the user to specify whether or not the location of the property in the image (e.g., blue at the bottom of the image or blue anywhere) is significant.
>
> Query by image similarity, wherein a user provides an entire image as a query target and the system finds images that are visually similar.

Jain, col. 4, lines 21-32. In this cited text, there is no disclosure or implication of subjective experiments performed with human observers, nor is there disclosure that perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers. Instead, Applicants read the cited text of Jain as allowing a user to provide visual queries, such as an image property or a provided image, which are then used to search a database (see, e.g., col. 4, lines 39-57 of Jain).

The Examiner cites FIG. 1A, elements 102, 104, and 112 (FIG. 1A of Jain is shown below):



APPLICATION MODULES

FIG. 1A

Jain states the following about element 112: "The visual input is provided to the VIR Engine 120. An 'Insertion' module 112 is used to provide one or more new images to be added to a database 132 accessible by the database engine 130. The new image(s) are provided as inputs to the VIR Engine 120." Jain, col. 9, lines 25-28. Jain also states the following:

> The image analysis module 122 receives inputs from either module 108 or 110 to generate a query target or from the insertion module 112 for adding a new image into the database 132. The output of the image analysis module 122 is a feature vector (FV) that describes the visual object passed to it by one of modules 108, 110 or 112. The FV is passed on to the database engine 130.

Jain, col. 9, lines 40-45. In this cited text and figure from Jain, there is no disclosure or implication of subjective experiments performed with human observers, nor is there disclosure that perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

The Examiner also cites col. 8, lines 32-35 of Jain. Jain states the following:

> The overall similarity between two images lies literally "in the eye of the beholder." In other words, the perceptual distance between images is not computable in terms of topological metrics. The same user will also change his or her interpretation of similarity depending on the task at hand. To express this subjective element, the VIR interface provides functions to allow the user to control which relative combinations of individual distances satisfies his or her needs. As the user changes the relative importance of primitives by adjusting a set of weighting factors (***at query time***), the VIR system incorporates the weight values into the similarity computation between feature vectors.

Jain, col. 8, lines 24-35, the first paragraph in a section entitled "Primitive Weighting" (emphasis added). It is noted that the adjustment is performed ***at query time***. Jain also states the following:

> The VIR Engine is a library-based tool kit that is delivered in binary form (an object library with header file interfaces) on various platforms, and provides an American National Standards Institute (ANSI) "C" language interface to the application developer. It provides access to the technology of Visual Information Retrieval (VIR), which allows images to be mathematically

> characterized and compared to one another on the basis of "visual similarity". Applications may now search for images or rank them based on "what they look like". The VIR Engine looks at the pixel data in the images, and analyzes the data with respect to visual attributes such as color, texture, shape, and structure. These visual attributes are called "primitives", and the image characterization is built up from these. Images which have been analyzed may then be compared mathematically to determine their similarity value or "score". Images are analyzed once, and the primitive data is then used for fast comparisons.

Jain, col. 6, lines 13-31. Applicants read this cited text as indicating that — at query time — the user in Jain can adjust a set of weighting factors that weight individual ones of the primitives, such as color, texture, shape, and structure. The VIR engine in Jain then uses the weight values when performing similarity computations between feature vectors (e.g., having values corresponding to the primitives).

This cited text from Jain does not disclose or imply that there are subjective experiments performed with human observers, nor does this disclose or imply that perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

The Examiner also cites col. 11, lines 43-59 of Jain:

> The user 102 (FIG. 1A) preferably initiates query generation 242 by either utilizing the query canvas 108 to create a query, or browses 110 the available file system to locate an existing object to use as the query, or browses 246 the database store 132 (FIG. 1A and FIG. 5B) to identify an image that has already been analyzed by the analysis module 122. In the last situation, if the image is already in the database 132, a feature vector has been computed and is retrieved at state 247 from a feature vector storage portion 264 of the database 132. A target image I.sub.T 248 results if either of the query canvas module 108 or browse file system module 110 are used to generate a query. The target image 248 is input to the analysis module 122 to generate a feature vector for the target image as the output. Because of the importance of the primitives in the system 100, a digression is now made to describe the base system primitives.

Jain, col. 11, lines 42-53. In this cited text, there is no disclosure or implication of subjective experiments performed with human observers, nor is there disclosure that perceptual features

and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

Finally, the Examiner cites col. 18, lines 7-26 of Jain:

> Two examples of utilizing weights with the primitives by use of the weights sliders (e.g., 208) in the query window 200 (FIG. 3) are as follows:
>
> Texture: The VIR Engine evaluates pattern variations within narrow sample regions to determine a texture value. It evaluates granularity, roughness, repetitiveness, and so on. Pictures with strong textural attributes--a sandstone background for example--tend to be hard to catalog with keywords. A visual search is the best way to locate images of these types. For best results, a user should set Texture high when the query image is a rough or grainy background image and low if the query image has a central subject in sharp focus or can be classified as animation or clip-art.
>
> Structure: The VIR Engine evaluates the boundary characteristics of distinct shapes to determine a structure value. It evaluates information from both organic (photographic) and vector sources (animation and clip art) and can extrapolate shapes partially obscured. Polka dots, for example, have a strong structural element. For best results, a user should set Structure high when the objects in the query image have clearly defined edges and low if the query image contains fuzzy shapes that gradually blend from one to another.

Jain, col. 18, lines 4-27. What this cited text appears to indicate is that a user can use sliders 208 to adjust weights associated with the primitives of Texture and Structure. In this cited, there is no disclosure or implication of subjective experiments performed with human observers, nor is there disclosure that perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

The Examiner asserts that the sections of Jain cited above clearly show that human observers/users/operators are heavily involved in the process of deriving perceptual features and category processing. Even if this assertion is true (which Applicants do not admit), there is no indication in Jain that such derivation of perceptual features and category processing involves subjective experiments performed with human observers, or that

perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

Applicants have reviewed all of Jain and cannot find disclosure or implication of subjective experiments performed with human observers, nor can Applicants find disclosure or implication that perceptual features and their combinations or perceptual semantic categories are derived at least in part through subjective experiments performed with human observers.

Consequently, independent claims 1 and 25 are patentable over Jain for at least the reasons given above, as claim 1 recites the subject matter of "wherein the perceptual features and their combinations are derived at least in part through subjective experiments performed with human observers" and claim 25 recites the subject matter of "where the perceptual features and their combinations are derived through subjective experiments performed with human observers". Similarly, claim 14 recites "a stored program for determining the semantic meaning of images in accordance with a set of perceptual semantic categories that were previously derived at least in part through subjective experiments performed with human observers and that represent important semantic cues in the human perception of images". The dependent claims are all allowable at least by virtue of their dependency from allowable independent claims. Thus, the individual merits of the dependent claims need not be discussed at this juncture.

Conclusion

Based on the foregoing arguments, it should be apparent that claims 1-28 are thus allowable over the reference(s) cited by the Examiner, and the Examiner is respectfully requested to reconsider and remove the rejections. The Examiner is invited to call the undersigned attorney for any issues.

S.N. 10/033,597
Art Unit: 2624

Respectfully submitted:

Robert J. Mauri, Reg. No.: 41,180

2/26/07
Date

Customer No.: 48237

HARRINGTON & SMITH, LLP
4 Research Drive
Shelton, CT 06484-6212

Telephone:	(203)925-9400
Facsimile:	(203)944-0245
email:	rmauri@hspatent.com

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. BOX 1450, Alexandria, VA 22313-1450.

Date

Name of Person Making Deposit